

FIG. 1

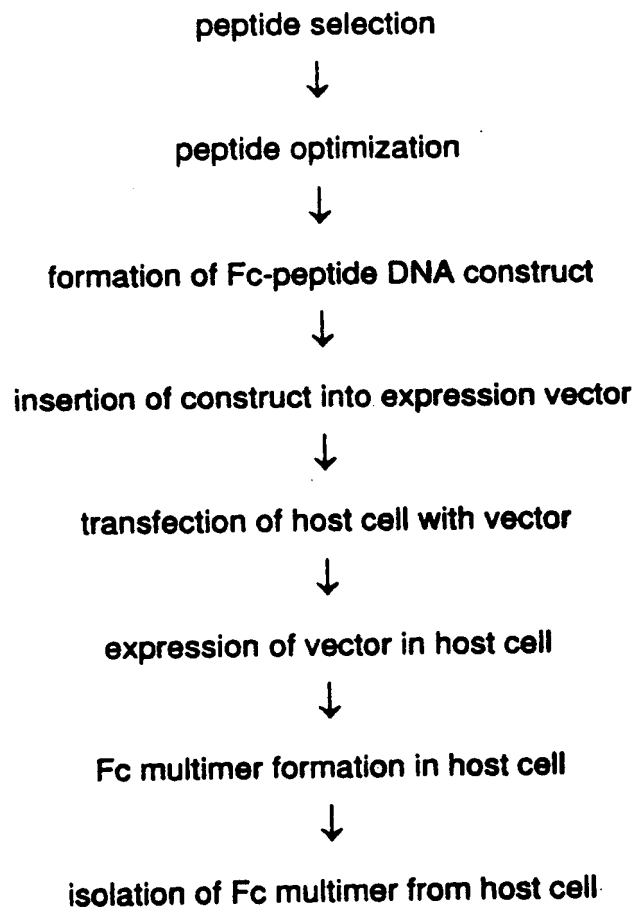


FIG. 2A

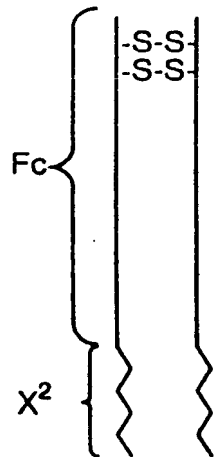


FIG. 2B

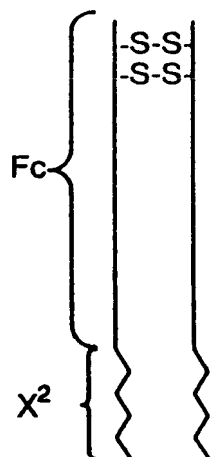


FIG. 2C

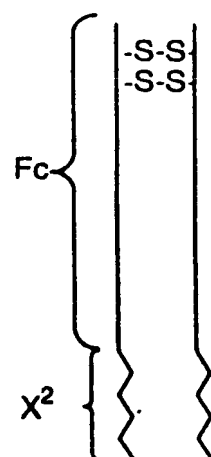


FIG. 2D

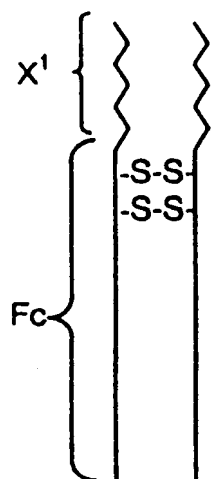


FIG. 2E

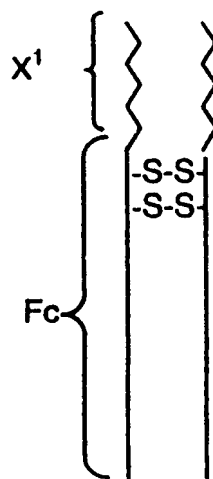


FIG. 2F

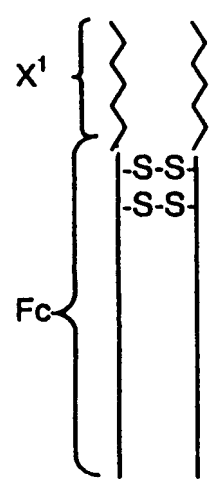


FIG. 3A

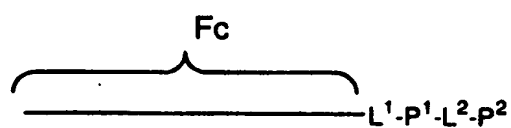


FIG. 3B

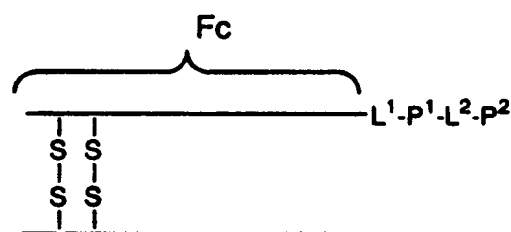


FIG. 3C

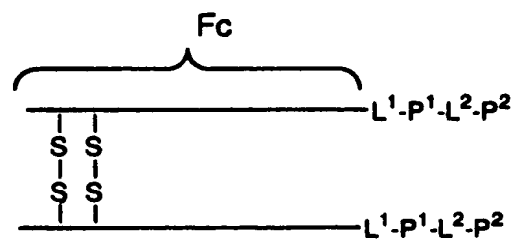


FIG. 4

```

ATGGACAAAAC TCACACATGTCCACCTTGTCCAGCTCCGGAAC TCCTGGGGGGACCGTCA
1  -----+-----+-----+-----+-----+-----+-----+-----+ 60
TACCTGTTTTGAGTGTGTACAGGTGGAACAGGTCGAGGCCTTGAGGACCCCCCTGGCAGT
a      M D K T H T C P P C P A P E L L G G P S .
GTCTTCCTCTTCCCCCAAAACCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTC
61  -----+-----+-----+-----+-----+-----+-----+-----+ 120
CAGAAGGAGAAGGGGGGTTTGGGTTCTGTGGGAGTACTAGAGGGCCTGGGGACTCCAG
a      V F L F P P K P K D T L M I S R T P E V .
ACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTG
121  -----+-----+-----+-----+-----+-----+-----+-----+ 180
TGTACGCACCACCACCTGCACTCGGTGCTTCTGGGACTCCAGTTCAAGTTGACCATGCAC
a      T C V V V D V S H E D P E V K F N W Y V .
GACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCAGG
181  -----+-----+-----+-----+-----+-----+-----+-----+ 240
CTGCCGCACCTCCACGTATTACGGTCTGTTCGGCGCCCTCCTCGTCATGTTGTCGTGC
a      D G V E V H N A K T K P R E E Q Y N S T .
TACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTAC
241  -----+-----+-----+-----+-----+-----+-----+-----+ 300
ATGGCACACCAGTCGCAGGAGTGGCAGGACGTGGTCTGACCGACTTACCGTTCCTCATG
a      Y R V V S V L T V L H Q D W L N G K E Y .
AAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCATCGAGAAAACCATCTCCAAGCC
301  -----+-----+-----+-----+-----+-----+-----+-----+ 360
TTCACGTTCCAGAGGTGTTTCGGGAGGGTGGGGGTAGCTCTTTTGGTAGAGGTTTCGG
a      K C K V S N K A L P A P I E K T I S K A .
AAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACC
361  -----+-----+-----+-----+-----+-----+-----+-----+ 420
TTTCCCGTCGGGGCTCTTGGTGTCCACATGTGGGACGGGGGTAGGGCCCTACTCGACTGG
a      K G Q P R E P Q V Y T L P P S R D E L T .
AAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTG
421  -----+-----+-----+-----+-----+-----+-----+-----+ 480
TTCTTGGTCCAGTCGGACTGGACGGACCAGTTTCCGAAGATAGGGTCGCTGTAGCGGCAC
a      K N Q V S L T C L V K G F Y P S D I A V .
GAGTGGGAGAGCAATGGGCAGCCGGAACAAC TACAAGACCACGCCTCCCGTGCTGGAC
481  -----+-----+-----+-----+-----+-----+-----+-----+ 540
CTCACCCCTCTCGTTACCCGTCGGCCTCTTGTGATGTTCTGGTGGGAGGGCACGACCTG
a      E W E S N G Q P E N N Y K T T P P V L D .
TCCGACGGCTCCTTCTTCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAG
541  -----+-----+-----+-----+-----+-----+-----+-----+ 600
AGGCTGCCGAGGAAGAAGGAGATGTCGTTGAGTGGCACCTGTTCTCGTCCACCGTCGTC
a      S D G S F F L Y S K L T V D K S R W Q Q .
GGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAG
601  -----+-----+-----+-----+-----+-----+-----+-----+ 660
CCCTTGCAAGAGTACGAGGCACTACGTACTCCGAGACGTGTTGGTGATGTGCGTCTTC
a      G N V F S C S V M H E A L H N H Y T Q K .
AGCCTCTCCCTGTCTCCGGGTAAA
661  -----+-----+-----+-----+-----+-----+-----+-----+ 684
TCGGAGAGGGACAGAGGCCCATTT

```

FIG. 5

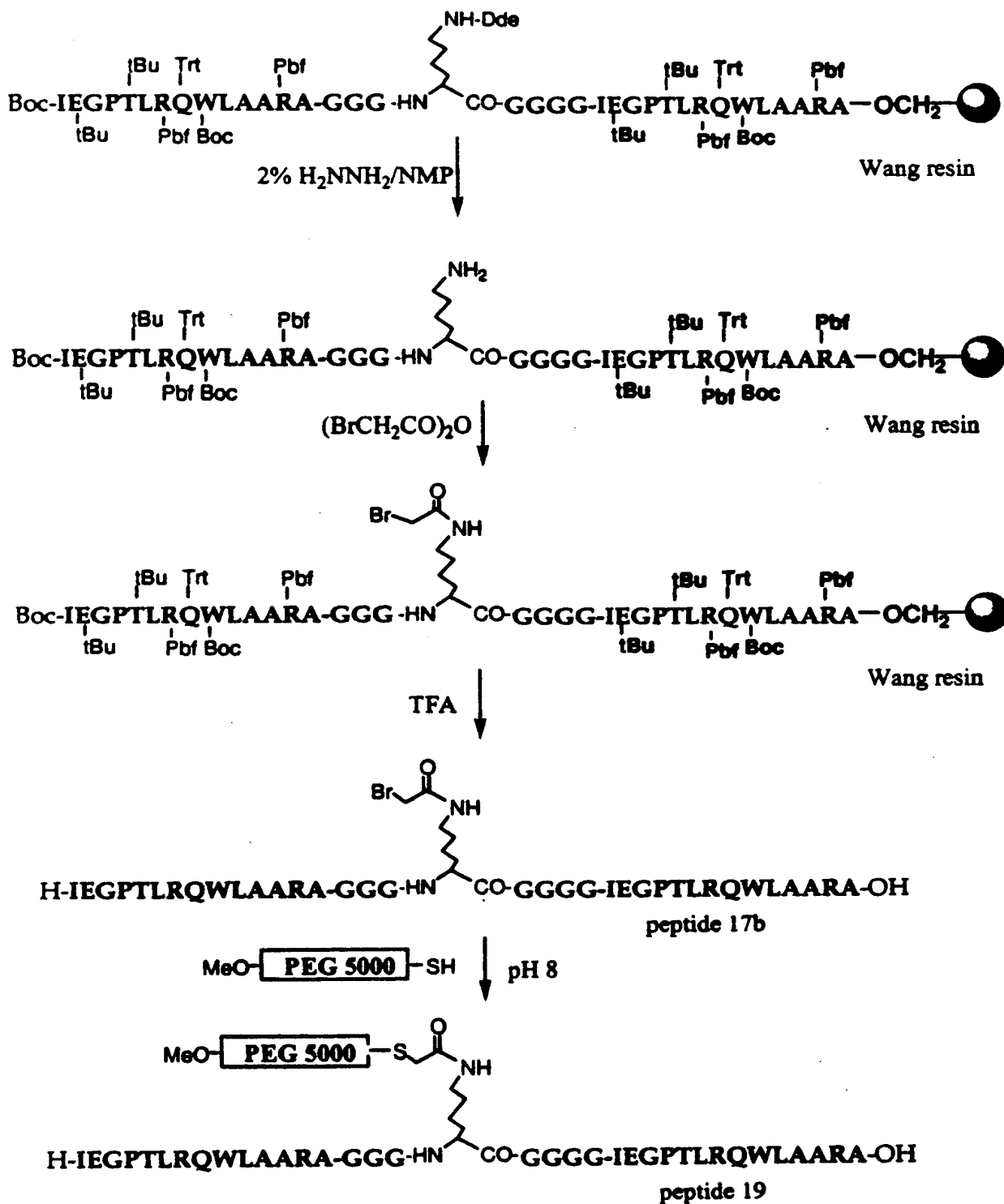


FIG. 6

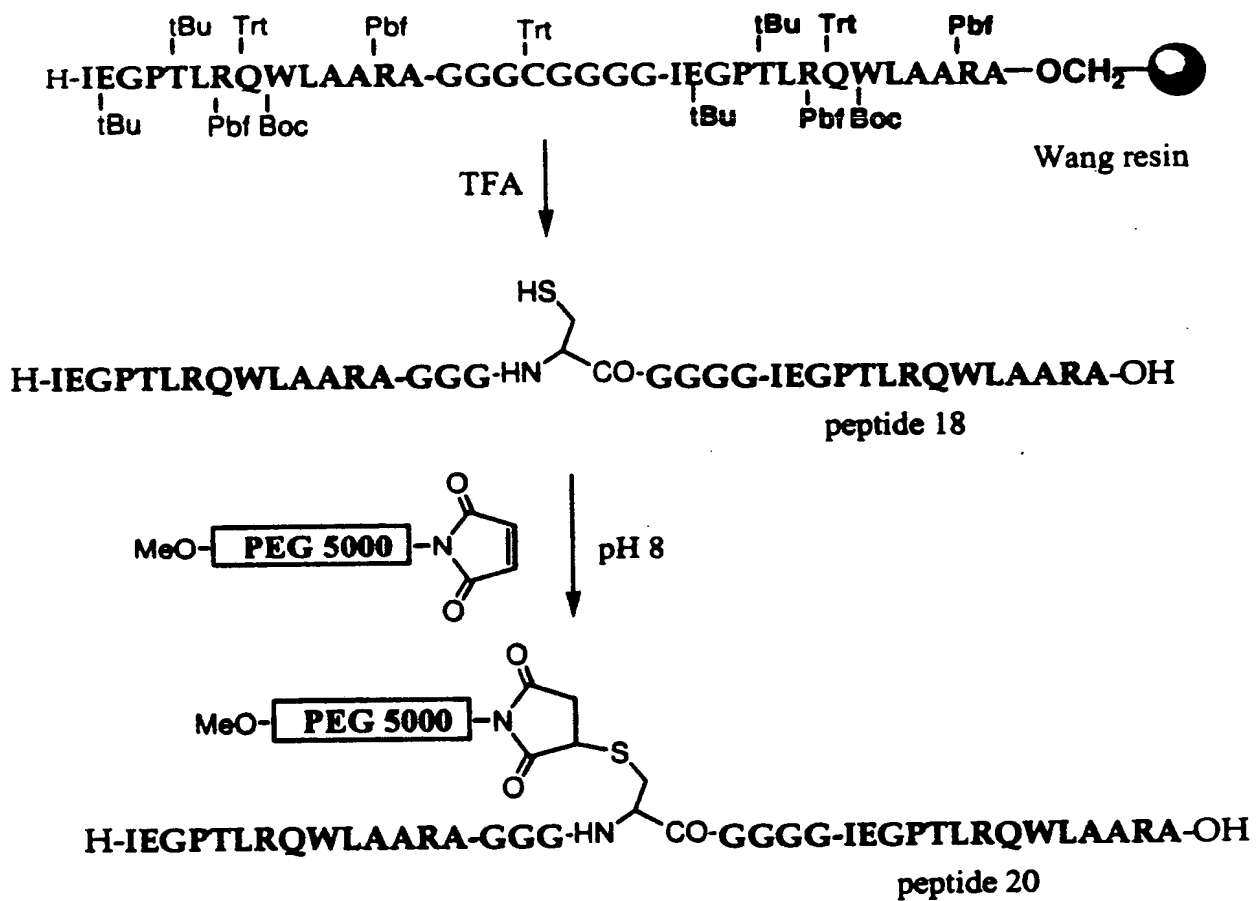


FIG. 7

XbaI
|
1 TCTAGATTTGTTTAACTAATTAAAGGAGGAATAACATATGGACAAAACACACATGTC
-----+-----+-----+-----+-----+ 60
AGATCTAAACAAAATTGATTAATTCCTCCTTATTGTATACCTGTTTTGAGTGTGTACAG
M D K T H T C P -
c
CACCTTGTCAGCTCCGGAACCTCTGGGGGGACCGTCAGTCTTCCTCTTCCCCCAAAC
-----+-----+-----+-----+-----+ 120
61 GTGGAACAGGTCGAGGCCTTGAGGACCCCTGGCAGTCAGAAGGAGAAGGGGGTTTGT
P C P A P E L L G G P S V F L F P P K P -
c
CCAAGGACACCCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGACGTGA
-----+-----+-----+-----+-----+ 180
121 GGTTCCTGTGGGAGTACTAGAGGGCCTGGGGACTCCAGTGTACGCACCACCACCTGCACT
K D T L M I S R T P E V T C V V V D V S -
c
GCCACGAAGACCCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATG
-----+-----+-----+-----+-----+ 240
181 CGGTGCTTCTGGGACTCCAGTTCAAGTTGACCATGCACCTGCCGCACCTCCACGTATTAC
H E D P E V K F N W Y V D G V E V H N A -
c
CCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGTGTGGTCAGCGTCTCTCA
-----+-----+-----+-----+-----+ 300
241 GGTTCCTGTTTCGGCGCCCTCCTCGTCATGTTGTCTGTCATGGCACACCAGTCCGAGGAGT
K T K P R E E Q Y N S T Y R V V S V L T -
c
CCGTCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAG
-----+-----+-----+-----+-----+ 360
301 GGCAGGACGTGGTCCTGACCGACTTACCGTTCTCATGTTTACGTTCCAGAGGTTGTTTC
V L H Q D W L N G K E Y K C K V S N K A -
c
CCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCAC
-----+-----+-----+-----+-----+ 420
361 GGGAGGGTCGGGGGTAGCTCTTTTGGTAGAGGTTTCGGTTTCCCGTCGGGGCTCTTGGTG
L P A P I E K T I S K A K G Q P R E P Q -
c
AGGTGTACACCCTGCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTCAGCCTGACCT
-----+-----+-----+-----+-----+ 480
421 TCCACATGTGGGACGGGGTAGGGCCCTACTCGACTGGTTCTTGGTCCAGTCCGACTGGA
V Y T L P P S R D E L T K N Q V S L T C -
c
GCCTGGTCAAAGGCTTCTATCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGC
-----+-----+-----+-----+-----+ 540
481 CGGACCAGTTTCCGAAGATAGGGTCGCTGTAGCGGCACCTCACCTCTCGTTACCCGTCG
L V K G F Y P S D I A V E W E S N G Q P -
c
CGGAGAACAACCTACAAGACCACGCCTCCCGTGTGGACTCCGACGGCTCTTCTTCTCTCT
-----+-----+-----+-----+-----+ 600
541 GCCTCTTGTGATGTTCTGCTGCGGAGGGCAGACCTGAGGCTGCCGAGGAAGAAGGAGA
E N N Y K T T P P V L D S D G S F F L Y -
c
ACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCG
-----+-----+-----+-----+-----+ 660
601 TGTCTGTTTCGAGTGGCACCTGTTCTCGTCCACCGTCGTCCCTTGCAGAAGAGTACGAGGC
S K L T V D K S R W Q Q G N V F S C S V -
c
TGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTA
-----+-----+-----+-----+-----+ 720
661 ACTACGTACTCCGAGACGTGTTGGTGATGTGCGTCTTCTCGGAGAGGGACAGAGGCCAT
M H E A L H N H Y T Q K S L S L S P G K -
c
AAGGTGGAGGTGGTGGTATCGAAGGTCCGACTCTGCGTCAGTGGCTGGCTGCTCGTGCTT
-----+-----+-----+-----+-----+ 780
721 TTCCACCTCCACCACCATAGCTTCCAGGCTGAGACGCAGTCACCGACCGACGAGCAGAA
G G G G G I E G P T L R Q W L A A R A * -
c
BamHI
|
AATCTCGAGGATCC
781 -----+----- 794
TTAGAGCTCCTAGG

FIG. 8

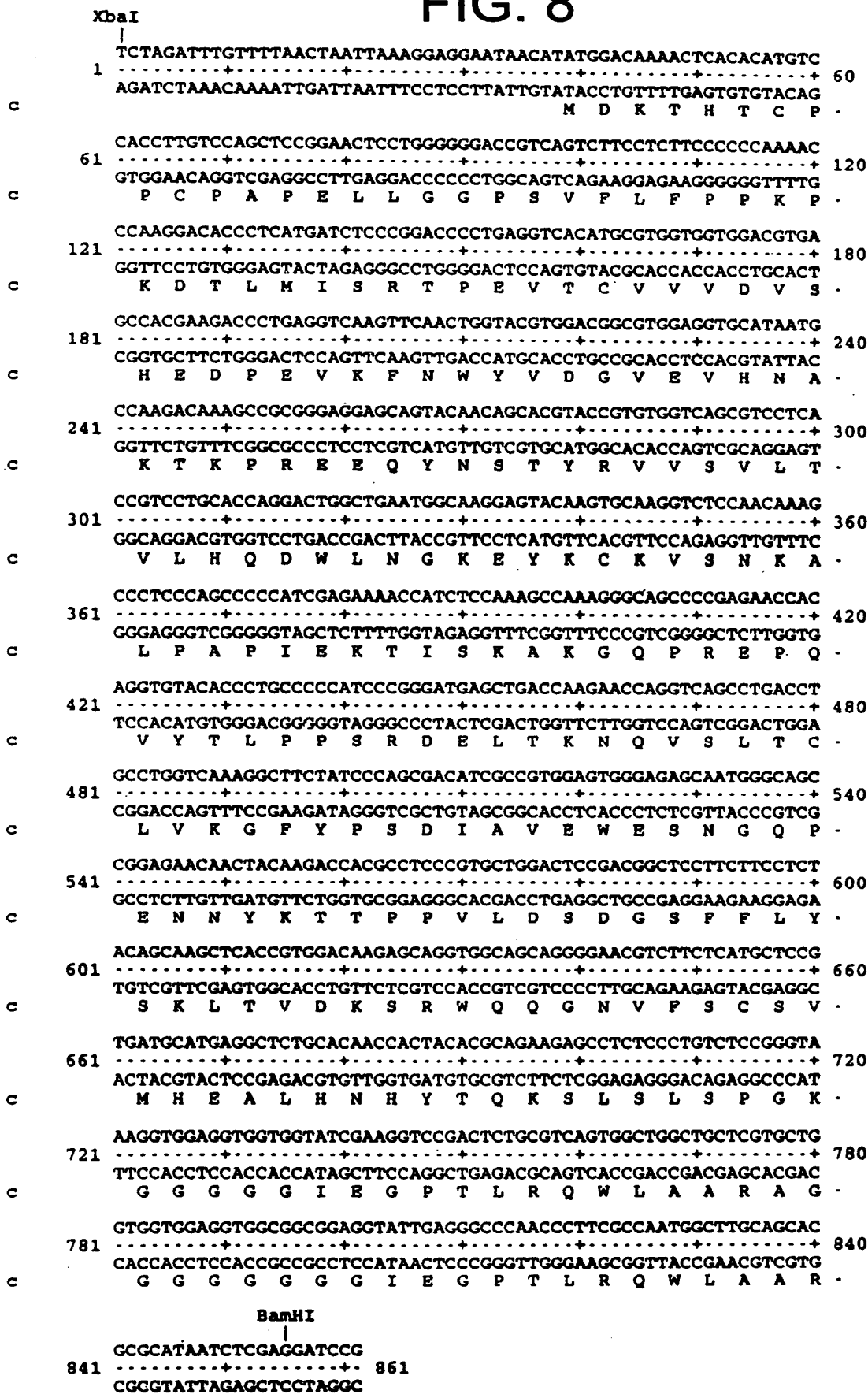


FIG. 9

[illegible]

FIG. 10

XbaI
|
1 TCTAGATTGTTTAACTAATTAAAGGAGGAATAACATATGATCGAAGGTCCGACTCTGC 60
AGATCTAAACAAAATTGATTAATTTCTCCTTATTGTATACTAGCTTCCAGGCTGAGACG
M I E G P T L R -

61 GTCAGTGGCTGGCTGCTCGTGCTGGTGGAGGCGGTGGGGACAAACTCACACATGTCCAC 120
CAGTCACCGACCGACGAGCAGCACCACCTCCGCCACCCCTGTTTGTAGTGTGTACAGGTG
Q W L A A R A G G G G G D K T H T C P P -

121 CTTGCCCAGCACCTGAACTCCTGGGGGACCGTCAGTTTTCTCTTCCCCCAAAACCCA 180
GAACGGGTCGTGGACTTGAGGACCCCTGGCAGTCAAAGGAGAAGGGGGTGGTGGT
C P A P E L L G G P S V F L F P P K P K -

181 AGGACACCCTCATGATCTCCCGACCCCTGAGGTACATGCGTGGTGGTGGACGTGAGCC 240
TCCTGTGGGAGTACTAGAGGGCCTGGGGACTCCAGTGTACGCACCACCACCTGCACTCGG
D T L M I S R T P E V T C V V V D V S H -

241 ACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCA 300
TGCTTCTGGGACTCCAGTTCAAGTTGACCATGCACCTGCCGCACCTCCACGTATTACGGT
E D P E V K F N W Y V D G V E V H N A K -

301 AGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGTGTGGTCCAGCGTCTCACCG 360
TCTGTTTCGGCGCCCTCCTCGTCATGTTGTCGTGCATGGCACACCAGTCGCAGGAGTGGC
T K P R E E Q Y N S T Y R V V S V L T V -

361 TCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCC 420
AGGACGTGGTCTTGACCGACTTACCGTTCTCATGTTTACGTTCCAGAGGTTGTTTCGGG
L H Q D W L N G K E Y K C K V S N K A L -

421 TCCCAGCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGG 480
AGGGTCGGGGGTAGCTCTTTTGGTAGAGTTTCGGTTTCCCGTCGGGGCTCTTGGTGTCC
P A P I E K T I S K A K G Q P R E P Q V -

481 TGTACACCCTGCCCCATCCCGGATGAGCTGACCAAGAACCAGGTACGCTGACCTGCC 540
ACATGTGGGACGGGGGTAGGGCCCTACTCGACTGGTCTTGGTCCAGTCGGACTGGACGG
Y T L P P S R D E L T K N Q V S L T C L -

541 TGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGG 600
ACCAGTTTCCGAAGATAGGGTCGCTGTAGCGGCACCTCACCTCTCGTTACCCGTCCGGC
V K G F Y P S D I A V E W E S N G Q P E -

601 AGAACAACCTACAAGACCACGCCTCCCGTGTGGACTCCGACGGCTCCTTCTCTCTTACA 660
TCTTGTGATGTTCTGGTGGGAGGGCAGCAGCTGAGGCTGCCGAGGAAGAAGGAGATGT
N N Y K T T P P V L D S D G S F F L Y S -

661 GCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGA 720
CGTTCGAGTGGCACCTGTTCTCGTCCACCGTCGTCCTTTCGAGAAGAGTACGAGGCACT
K L T V D K S R W Q Q G N V P S C S V M -

721 TGCATGAGGCTCTGCACAACCACTACACGAGAAGAGCCTCTCCCTGTCTCCGGGTAAAT 780
ACGTACTCCGAGACGTGTTGGTGTATGTGCGTCTTCTCGGAGAGGGACAGAGGCCCATTTA
H E A L H N H Y T Q K S L S L S P G K * -

BamHI
|
781 AATGGATCC 789
TTACCTAGG

FIG.11

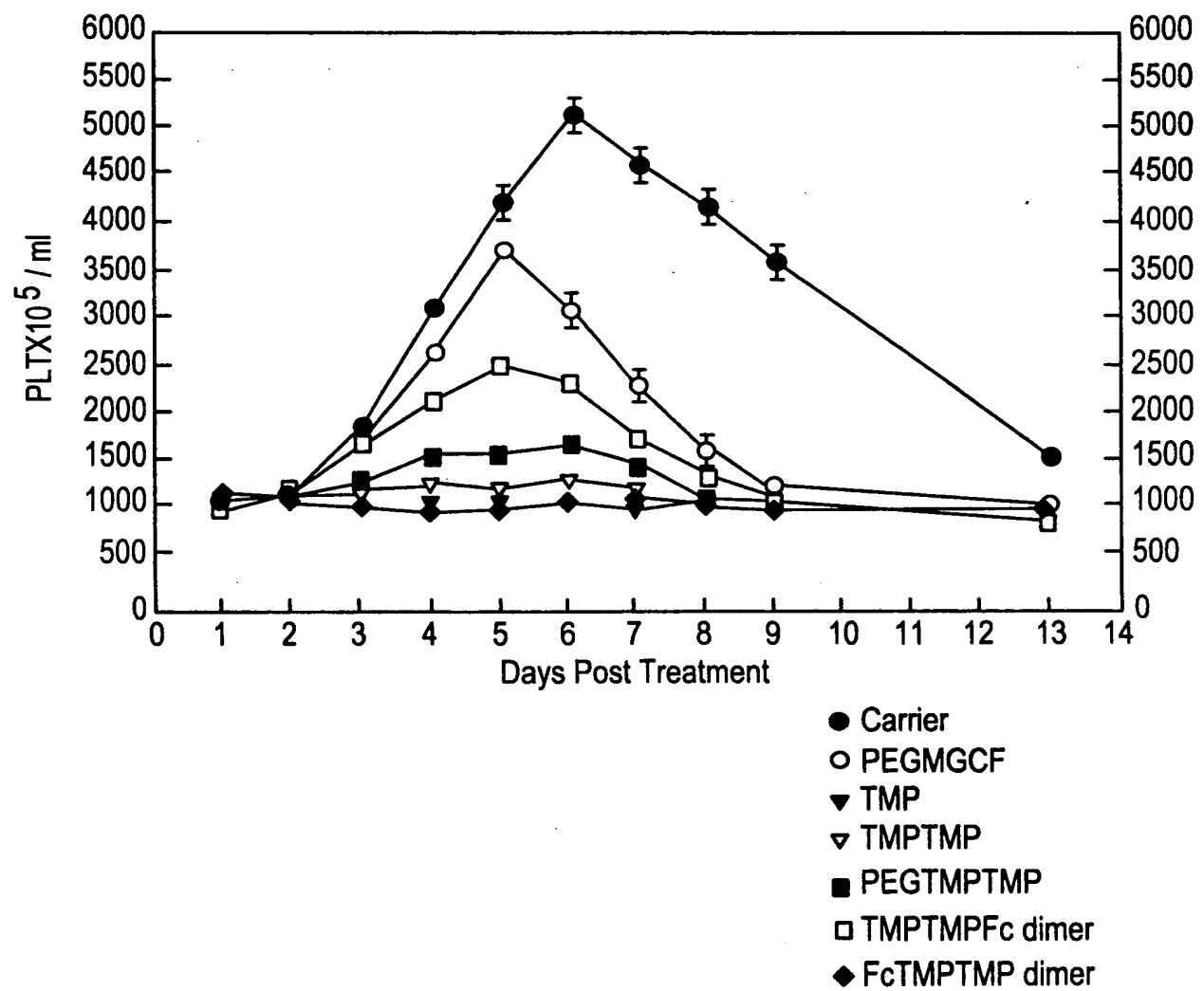


FIG.12

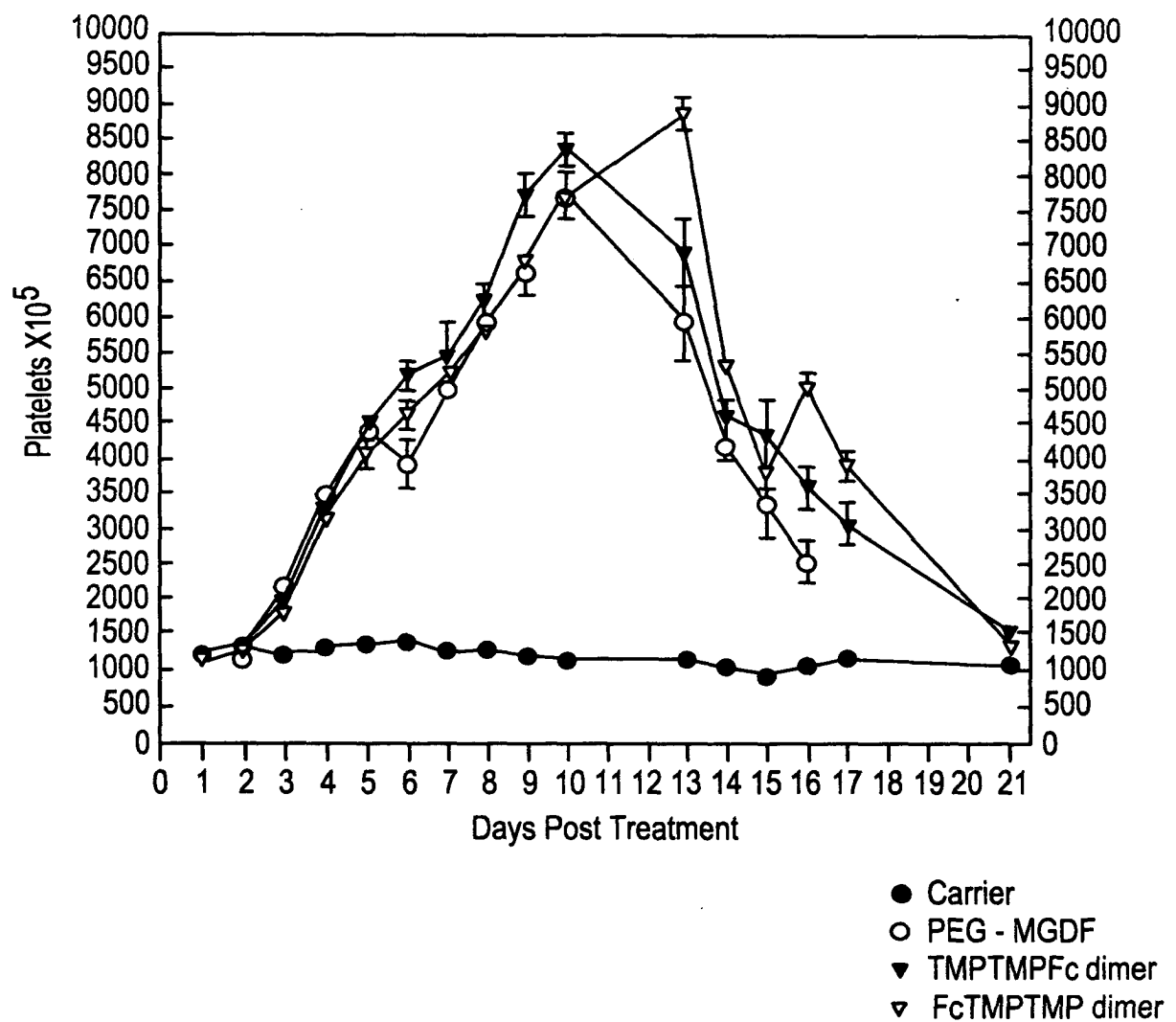


FIG. 13

XbaI
|

1 TCTAGATTTGTTTAACTAATTAAGGAGGAATAACATATGGACAAAACCTCACACATGTC 60
AGATCTAAACAAAATTGATTAATTTCTCTCTTATTGTATACCTGTTTTGAGTGTGTACAG
CACCTTGTCCAGCTCCGGAACCTCTGGGGGACCGTCAGTCTTCTCTTCCCCCAAAAC
GTGGAACAGGTCGAGGCCTTGAGGACCCCTGGCAGTCAGAAGGAGAAGGGGGTTTTG
CCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACGTGA
GGTTCCTGTGGGAGTACTAGAGGGCCTGGGGACTCCAGTGTACGCACCACCCTGCACT
GCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATG
CGGTGCTTCTGGGACTCCAGTTCAAGTTGACCATGCACCTGCCGCACCTCCACGTATTAC
CCAAGACAAAGCCCGGGAGGAGCAGTACAACAGCACGTACCGTGTGGTCAGCGTCCTCA
GGTTCCTGTTTCGGCGCCCTCCTCGTCATGTTGTCGTGCATGGCACACCAGTCGAGGAGT
CCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGAAGGTCTCCAACAAAG
GGCAGGACGTGGTCTTGACCGACTTACCGTTCCTCATGTTTACGTTCCAGAGGTTGTTTC
CCCTCCCAGCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACCAC
GGGAGGGTCTGGGGGTAGCTCTTTTGGTAGAGGTTTCGGTTTCCCGTCCGGGCTCTTGGTG
AGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGGTACAGCTGACCT
TCCACATGTGGGACGGGGTAGGGCCCTACTCGACTGGTCTTGGTCCAGTCCGACTGGA
GCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGACG
CGGACCAGTTTCCGAAGATAGGGTCGCTGTAGCGGCACCTCACCTCTCGTTACCCGTCG
CGGAGAACAACCTACAAGACCACGCCTCCCGTGTCTGGACTCCGACGGCTCCTTCTTCTCT
GCCTCTTGTGATGTTCTGGTGGGAGGGCAGACCTGAGGCTGCCGAGGAAGAAGGAGA
ACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCG
TGTCGTTTCGAGTGGCACCTGTTCTCGTCCACCGTCTGCCCTTGCAGAAGAGTACGAGGC
TGATGCATGAGGCTCTGCACAACCCTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTA
ACTACGTACTCCGAGACGTGTTGGTGTATGTGCGTCTTCTCGGAGAGGGACAGAGGCCCAT
AAGGTGGAGGTGGTGGTGGAGGTACTTACTCTTGCCACTTCGGCCCCGTGACTTGGGTTT
TTCCACCTCCACCACCACCTCCATGAATGAGAACGGTGAAGCCGGGCGACTGAACCCAAA
G G G G G G G T Y S C H F G P L T W V C
BamHI
|

781 GCAAACCGCAGGGTGGTTAATCTCGTGGATCC 812
CGTTTGGCGTCCCACCAATTAGAGCACCTAGG
K P Q G G *

FIG. 14

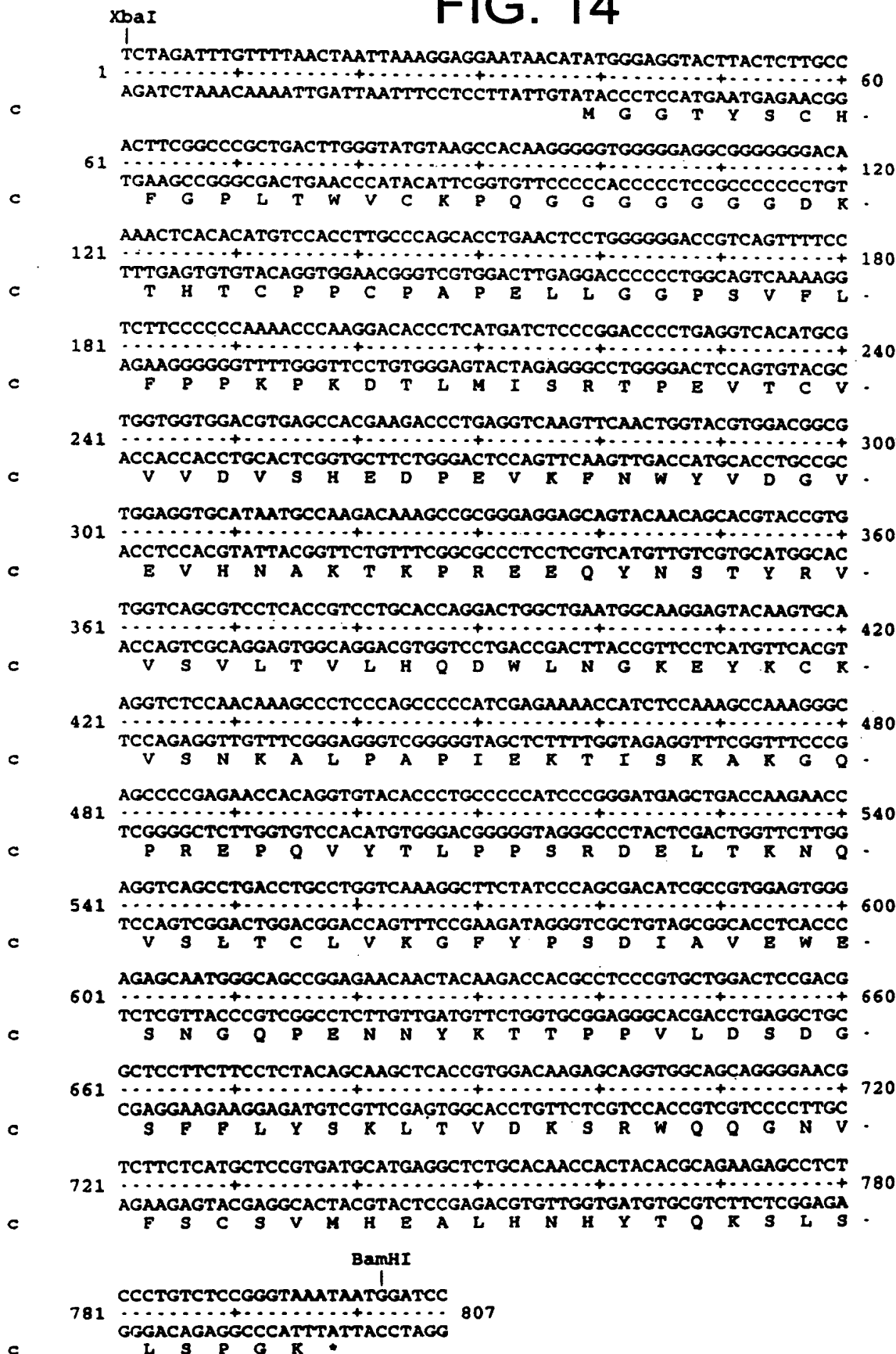


FIG. 15

XbaI
|

1 TCTAGATTGAGTTTAACTTTTAGAAGGAGGAATAAAATATGGGAGGTACTTACTCTTG 60
AGATCTAAACTCAAAATTGAAAATCTCCTCCTTATTTTATACCCTCCATGAATGAGAAC
M G G T Y S C

61 CCACTTCGGCCCACTGACTTGGGTTTGCAAACCGCAGGGTGGCGGCGGCGGCGGCGGTGG 120
GGTGAAGCCCGGTGACTGAACCCAAACGTTTGGCGTCCACCGCGCGCGCGCGCGCCACC
H F G P L T W V C K P Q G G G G G G G G

121 TACCTATTCCTGTCATTTTGGCCCGCTGACCTGGGTATGTAAGCCACAAGGGGTGGGGG 180
ATGGATAAGGACAGTAAACCGGGCGACTGGACCCATACATTGGGTGTTCCCCACCCCC
T Y S C H F G P L T W V C K P Q G G G G

181 AGGCGGGGGGACAAAACCTCACACATGTCCACCTTGCCCAGCACCTGAACTCCTGGGGGG 240
TCCGCCCCCTGTTTGGAGTGTGTACAGGTGGAACGGGTCGTGGACTTGAGGACCCCC
G G G D K T H T C P P C P A P E L L G G

241 ACCGTCAGTTTCTCTTCCCCCAAAACCAAGGACACCCCTCATGATCTCCCGGACCCC 300
TGGCAGTCAAAGGAGAAGGGGGTTTTGGGTTCTGTGGGAGTACTAGAGGGCCTGGGG
P S V F L F P P K P K D T L M I S R T P

301 TGAGGTCACATGCGTGGTGGTGACGTGAGCCACGAAGACCCCTGAGGTCAAGTTCAACTG 360
ACTCCAGTGTACGCACCACCACCTGCACTCGGTGCTTCTGGGACTCCAGTTCAAGTTGAC
E V T C V V V D V S H E D P E V K F N W

361 GTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGAGGAGCAGTACAA 420
CATGCACCTGCCGCACCTCCACGTATTACGGTTCTGTTTCGGCGCCCTCCTCGTCATGTT
Y V D G V E V H N A K T K P R E E Q Y N

421 CAGCACGTACCGTGTGGTCAGCGTCTCACCGTCTGCACCAGGACTGGCTGAATGGCAA 480
GTCGTGCATGGCACACCAAGTCGAGGAGTGGCAGGACGTGGTCTGACCGACTTACCGTT
S T Y R V V S V L T V L H Q D W L N G K

481 GGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCATCGAGAAAACCATCTC 540
CCTCATGTTACGTTCCAGAGGTTGTTTCGGGAGGGTCGGGGTAGCTCTTTTGGTAGAG
E Y K C K V S N K A L P A P I E K T I S

541 CAAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGATGA 600
GTTTCGGTTTCCCGTCCGGGCTCTTGGTGTCCACATGTGGGACGGGGTAGGGCCCTACT
K A K G Q P R E P Q V Y T L P P S R D E

601 GCTGACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCAGCGACAT 660
CGACTGGTCTTGGTCCAGTCGGACTGGACGGACAGTTTCCGAAGATAGGGTCGCTGTA
L T K N Q V S L T C L V K G F Y P S D I

661 CGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACCTACAAGACCACGCTCCCGT 720
GCGGCACCTCACCTCTCGTTACCCGTCCGCCTCTTGTGATGTTCTGGTGGGAGGGCA
A V E W E S N G Q P E N N Y K T T P P V

721 GCTGGACTCCGACGGCTCCTTCTCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTG 780
CGACCTGAGGCTGCCGAGGAAGAAGGAGATGTCGTTTCGAGTGGCACCTGTTCTCGTCCAC
L D S D G S F F L Y S K L T V D K S R W

781 GCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACAC 840
CGTCGTCCCTTGCAGAAGAGTACGAGGCACTACGTACTCCGAGACGTGTTGGTGATGTG
Q Q G N V F S C S V M H E A L H N H Y T

BamHI
|

841 GCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATAATGGATCC 881
CGTCTTCTCGGAGAGGGACAGAGGCCATTTATTACCTAGG
Q K S L S L S P G K

FIG. 16

XbaI
|
1 TCTAGATTGTTTTAACTAATTAAGGAGGAATAACATATGGACAAAACCTCACACATGTC 60
AGATCTAAACAAAATTGATTAATTTCTCTCTTATGTATACCTGTTTTGAGTGTGTACAG
M D K T H T C P .
CACCTTGCCCAGCACCTGAACTCCTGGGGGGACCGTCAGTTTTCTCTCTCCCCCAAAAC
61 GTGGAACGGGTCGTGGACTTGAGGACCCCCCTGGCAGTCAAAAGGAGAAGGGGGTTTTG 120
P C P A P E L L G G P S V F L P P P K P .
CCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACGTGA
121 GGTTCCTGTGGGAGTACTAGAGGGCTGGGGACTCCAGTGTACGCACCACCCTGCACCT 180
K D T L M I S R T P E V T C V V V D V S .
GCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATG
181 CGGTGCTTCTGGGACTCCAGTTCAAGTTGACCATGCACCTGCCGCACCTCCACGTATTAC 240
H E D P E V K F N W Y V D G V E V H N A .
CCAAGACAAAGCCGCGGAGGAGCAGTACAACAGCACGTACCGTGTGGTCCAGCGTCTCA
241 GGTTCGTGTTTCGGCGCCCTCCTCGTCATGTTGTGCGTGCATGGCACACAGTCGAGGAGT 300
K T K P R E E Q Y N S T Y R V V S V L T .
CCGTCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTCAAGGTCTCCAACAAAG
301 GGCAGGACGTGGTCTGACCGACTTACCGTTCTCATGTTTCACGTTCCAGAGGTTGTTTC 360
V L H Q D W L N G K E Y K C K V S N K A .
CCCTCCCAGCCCCATCGAGAAAACCATCTCCAAGGCCAAAGGGCAGCCCCGAGAACCAC
361 GGGAGGGTCGGGGGTAGCTCTTTTGGTAGAGGTTTCGGTTTCCCGTCGGGGCTCTTGGTG 420
L P A P I E K T I S K A K G Q P R E P Q .
AGGTGTACACCCTGCCTCCATCCCGGATGAGCTGACCAAGAACCAGGTACGCTGACCT
421 TCCACATGTGGGACGGAGGTAGGGCCCTACTCGACTGGTCTTGGTCCAGTCGGACTGGA 480
V Y T L P P S R D E L T K N Q V S L T C .
GCCTGGTCAAAGGCTTCTATCCACCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGC
481 CGGACCAGTTTCCGAAGATAGGTCGCTGTAGCGGCACCTCACCTCTCGTTACCCGTCG 540
L V K G F Y P S D I A V E W E S N G Q P .
CGGAGAACAACACTACAAGACACGCTCCCGTGTGACTCCGACGGCTCCTTCTCTCTCT
541 GCCTCTTGTGATGTTCTGGTGGGAGGGCAGACCTGAGGCTGCCGAGGAAGAAGGAGA 600
E N N Y K T T P P V L D S D G S F P L Y .
ACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCG
601 TGTCGTTTCGAGTGGCACCTGTTCTCGTCCACCGTCGTCCCTTGCAGAAGAGTACGAGGC 660
S K L T V D K S R W Q Q G N V F S C S V .
TGATGCATGAGGCTCTGCACAACCACTACACGAGAAGAGCCTCTCCCTGTCTCCGGGTA
661 ACTACGTACTCCGAGACGTGTTGGTGATGTGCGTCTTCTCGGAGAGGGACAGAGGCCCAT 720
M H E A L H N H Y T Q K S L S L S P G R .
AAGGTGGAGGTGGTGGCGGAGGTACTTACTCTTGCCACTTCGCCCCACTGACTTGGGTTT
721 TTCCACCTCCACCACCGCTCCATGAATGAGAACGGTGAAGCCGGGTGACTGAACCCAAA 780
G G G G G G G T Y S C H F G P L T W V C .
GCAAACCGCAGGTGGCGGGCGGCGGCGGTGACCTATCTCTGTCAATTTGGCCCCG
781 CGTTTGGCGTCCACCGCCCGCCCGCCCGCCCATGGATAAGGACAGTAAAACCGGGCG 840
K P Q G G G G G G G G T Y S C H F G P L .
BamHI
|
841 TGACCTGGGTATGTAAGCCACAAGGGGGTTAATCTCGAGGATCC 884
ACTGGACCCATACATTCCGTGTTCCCCCAATTAGAGCTCCTAGG
T W V C K P Q G G .

FIG. 17A

[AatII sticky end]
(position #4358 in pAMG21)

5' GCGTAACGTATGCATGGTCTCC-
3' TGCACGCATTGCATACGTACCAGAGG-

-CCATGCGAGAGTAGGGAAC TGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACT-
-GGTACGCTCTCATCCCTTGACGGTCCGTAGTTTATTTTGCTTTCCGAGTCAGCTTTCTGA-
-GGGCCTTTCGTTTTATCTGTTGTTTGTGCGGTGAACGCTCTCCTGAGTAGGACAAATCCGC-
-CCCGGAAAGCAAAATAGACAACAAACAGCCACTTGCGAGAGGACTCATCCTGTTTAGGCG-
-CGGGAGCGGATTTGAACGTTGCGAAGCAACGGCCCGGAGGGTGGCGGGCAGGACGCCCGC-
-GCCCTCGCTAAACTTGCAACGCTTCGTTGCCGGGCTCCACCGCCCCGTCTGCGGGCG-
-CATAAACTGCCAGGCATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCGT-
-GTATTTGACGGTCCGTAGTTTAATTCGTCTTCCGGTAGGACTGCCTACCGGAAAAACGCA-
AatII
-TTCTACAACTCTTTTGTTTATTTTCTAAATACATTCAAATATGGACGTCGTACTTAAC-
-AAGATGTTTGAGAAAACAAATAAAAAGATTTATGTAAGTTTATACCTGCAGCATGAATTG-
-TTTTAAAGTATGGGCAATCAATTGCTCCTGTTAAATTGCTTTAGAAATACTTTGGCAGC-
-AAAATTTCATACCCGTTAGTTAACGAGGACAATTTTAACGAAATCTTTATGAAACCGTCG-
-GGTTTGTTGTATTGAGTTTCATTTGCGCATTGGTTAAATGGAAAGTGACCGTGCGCTTAC-
-CCAAACAACATAACTCAAAGTAAACGCGTAACCAATTTACCTTTCAGTGGCAGCGGAATG-
-TACAGCCTAATATTTTTGAAATATCCCAAGAGCTTTTTCTTCGATGCCACGCTAAAC-
-ATGTCGGATTATAAAAACCTTATAGGGTCTCGAAAAAGGAAGCGTACGGGTGCGATTTG-
-ATTCTTTTCTCTTTTGGTTAAATCGTTGTTTGATTTATTATTGCTATATTTATTTTTC-
-TAAGAAAAAGAGAAAACCAATTTAGCAACAACTAAATAATAAACGATATAAATAAAAAG-
-GATAATTATCAACTAGAGAAGGAACAATTAATGGTATGTTTCATACACGCATGTAAAAATA-
-CTATTAATAGTTGATCTCTTCTTGTTAATTACCATAACAAGTATGTGCGTACATTTTAT-
-AACTATCTATATAGTTGTCTTCTCTGAATGTGCAAACTAAGCATTCCGAAGCCATTAT-
-TTGATAGATATATCAACAGAAAGAGACTTACACGTTTTGATTCGTAAGGCTTCGGTAATA-
-TAGCAGTATGAATAGGGAACTAAACCCAGTGATAAGACCTGATGATTTGCTTCTTTAA-
-ATCGTCATACTTATCCCTTTGATTTGGGTCACTATTCTGGACTACTAAAGCGAAGAAAT-
-TTACATTTGGAGATTTTTTATTTACAGCATTGTTTTCAAATATATTCCAATTAATCGGTG-
-AATGTAAACCTCTAAAAATAAATGTCGTAACAAAAGTTTATATAAGGTTAATTAGCCAC-
-AATGATTGGAGTTAGAATAATCTACTATAGGATCATATTTTATTAAATTAGCGTCATCAT-
-TTACTAACCTCAATCTTATTAGATGATATCCTAGTATAAAATAATTTAATCGCAGTAGTA-
-AATATTGCCTCCATTTTTTAGGGTAATTATCCAGAATTGAAATATCAGATTTAACCATAG-
-TTATAACGGAGGTAAAAATCCCATTAATAGGTCTTAACCTTTATAGTCTAAATTGGTATC-
-AATGAGGATAAATGATCGCGAGTAAATAATATTCACAATGTACCATTTTAGTCATATCAG-
-TTACTCCTATTTACTAGCGCTCATTTATTATAAGTGTTACATGGTAAATCAGTATAGTC-
-ATAAGCATTGATTAATATCATTATTGCTTCTACAGGCTTTAATTTTATTAATTATTCTGT-
-TATTCGTAACATAATTATAGTAATAACGAAGATGTCCGAAATTAATAAATTAATAAGACA-
-AAGTGTCGTCGGCATTATATGTCCTTCATACCCATCTCTTTATCCTTACCTATTGTTTGTG-
-TTCACAGCAGCCGTAAATACAGAAAGTATGGGTAGAGAAATAGGAATGGATAACAAACAG-
-GCAAGTTTTGCGTGTTATATATCATTAAAACGGTAATAGATTGACATTTGATTCTAATAA-
-CGTTCAAACGCACAATATATAGTAATTTTGCCATTATCTAACTGTAACTAAGATTATT-

FIG. 17B

- ATTGGATTTTTGTCACACTATTATATCGCTTGAAATACAATTGTTTAAACATAAGTACCTG -
- TAACCTAAAAACAGTGTGATAATATAGCGAACTTTATGTAAACAAATTGTATTCATGGAC -

- TAGGATCGTACAGGTTTACGCAAGAAAATGGTTTGTTATAGTCGATTAATCGATTTGATT -
- ATCCTAGCATGTCCAAATGCGTTCCTTTACCAAACAATATCAGCTAATTAGCTAAACTAA -

- CTAGATTTGTTTTAACTAATTAAAGGAGGAATAACATATGGTTAACGCGTTGGAATTCGA -
- GATCTAAACAAAATTGATTAATTTCTCCTTATTGTATACCAATTGCGCAACCTTAAGCT -

- GCTCACTAGTGTGCGACCTGCAGGGTACCATGGAAGCTTACTCGAGGATCCGCGGAAAGAA -
- CGAGTGATCACAGCTGGACGTCCCATGGTACCTTCGAATGAGCTCCTAGGCGCCTTTCTT -

- GAAGAAGAAGAAGAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATA -
- CTTCTTCTTCTTCTTTCGGGCTTTCCTTCGACTCAACCGACGACGGTGGCGACTCGTTAT -

- ACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTTGCTGAAAGGAGG -
- TGATCGTATTGGGGAACCCCGGAGATTTGCCCAGAACTCCCCAAAAACGACTTTCCTCC -

- AACCGCTCTTCACGCTCTTCACGC 3' [SacII sticky end]
- TTGGCGAGAAGTGCGAGAAGTG 5' (position #5904 in pAMG21)

FIG.18A - 1

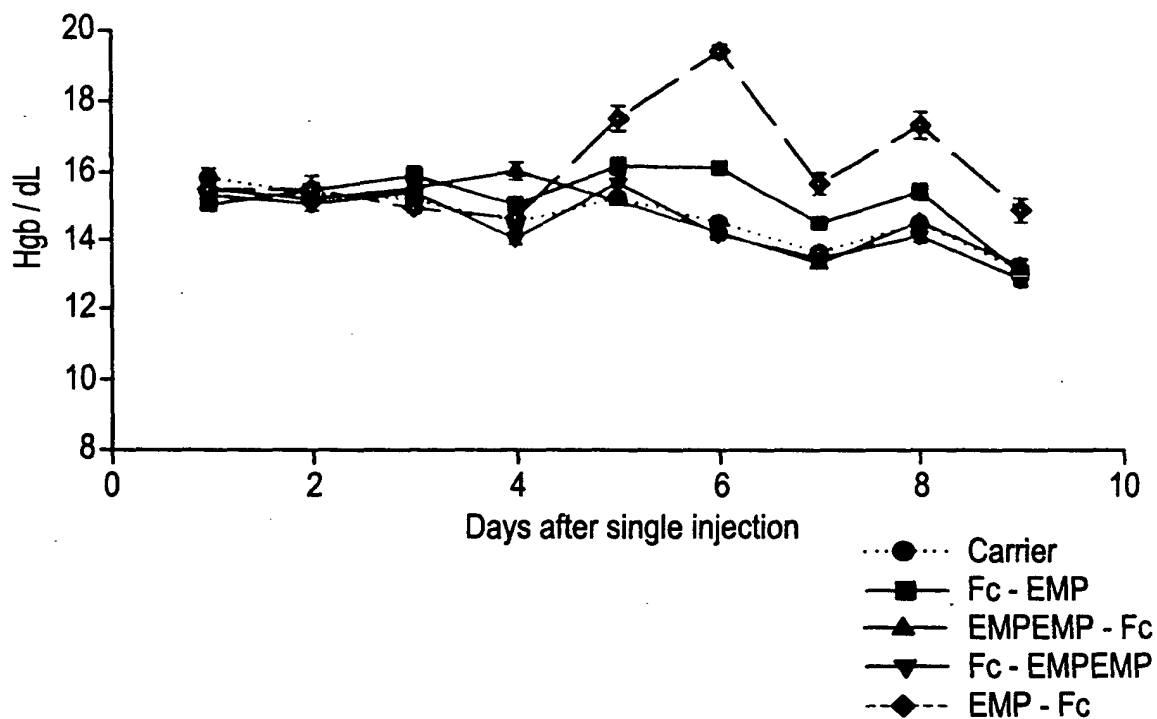


FIG.18A - 2

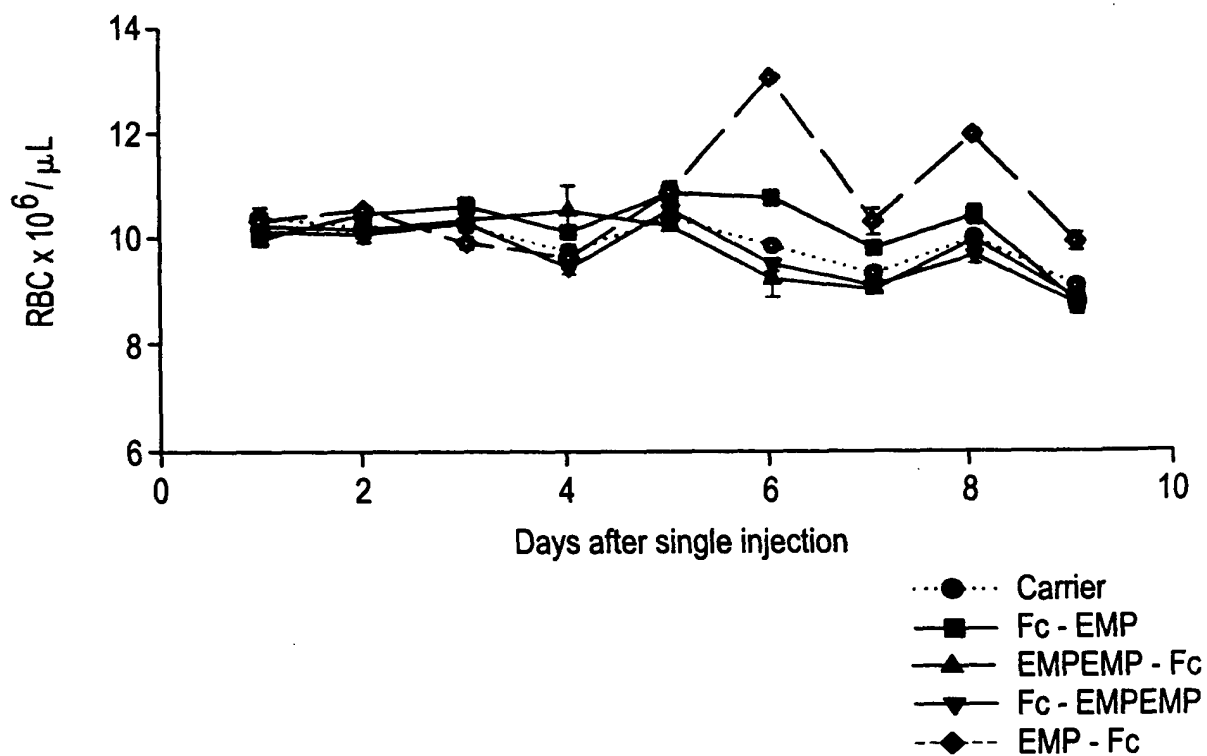


FIG.18A - 3

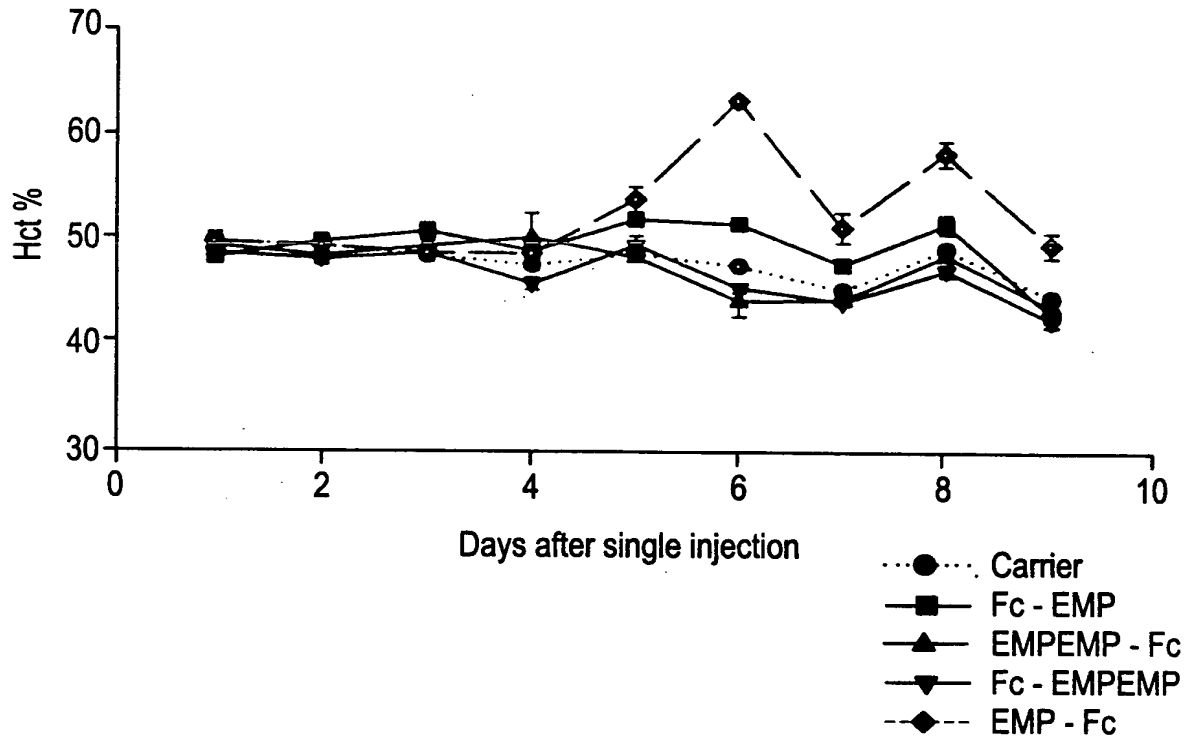


FIG.18B - 1

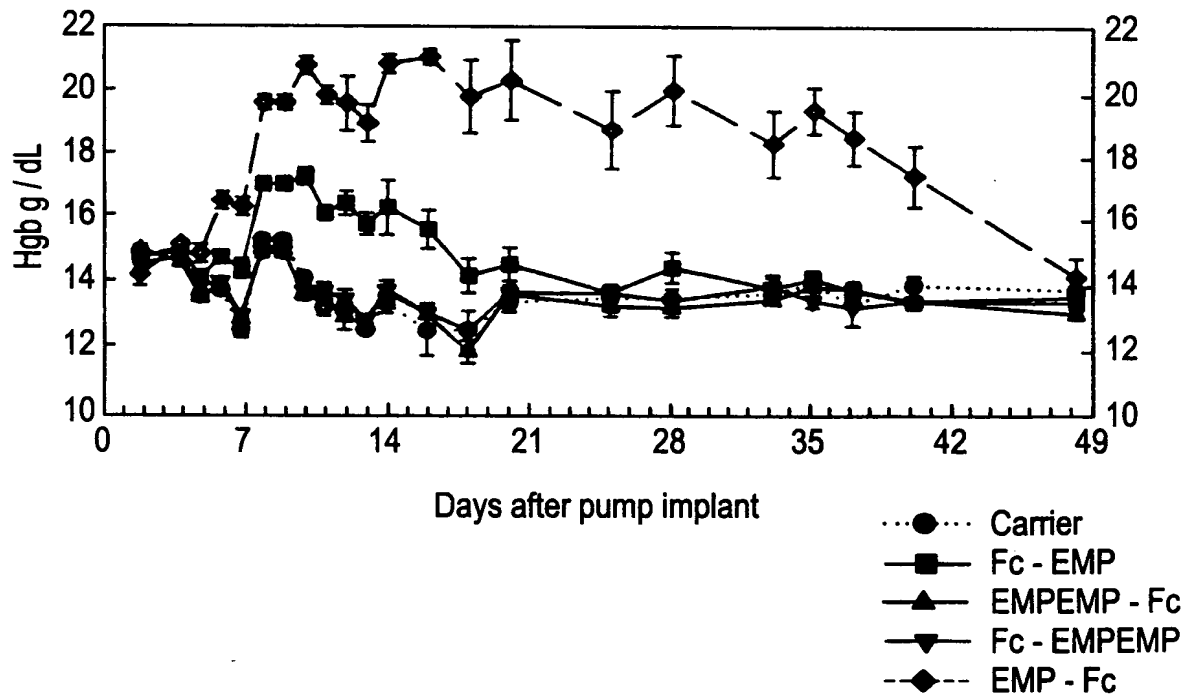


FIG.18B - 2

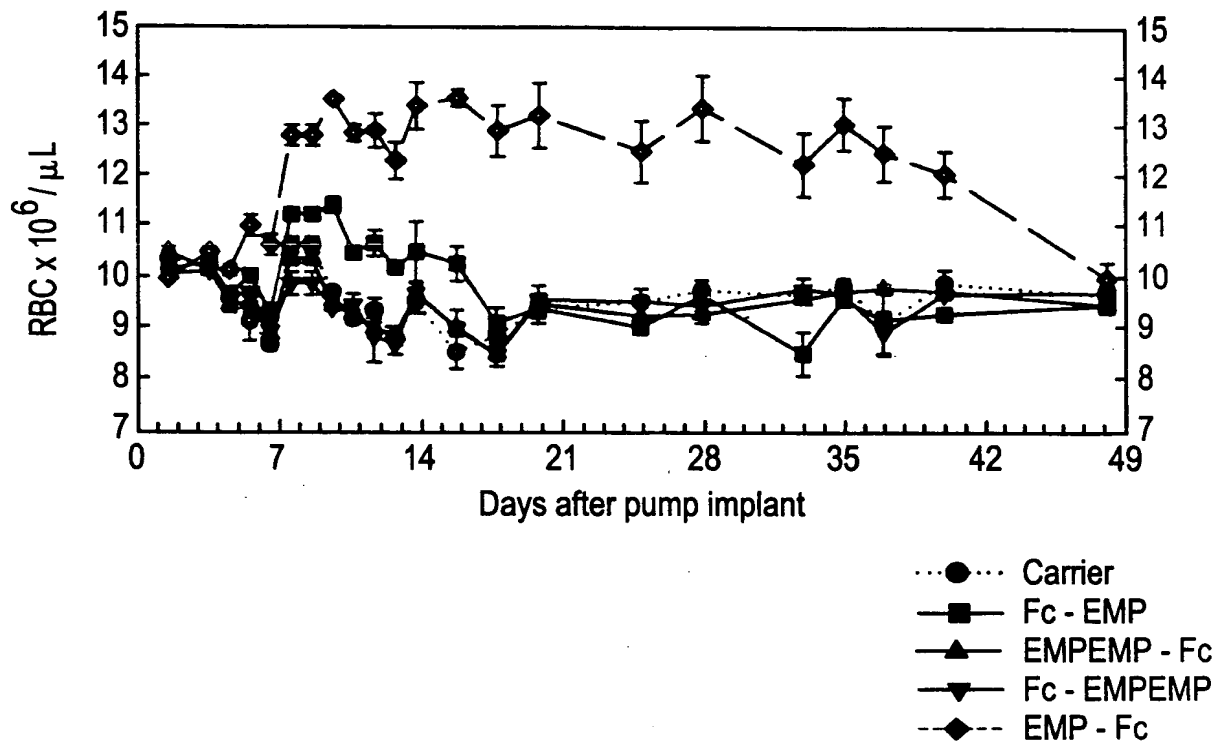


FIG.18B - 3

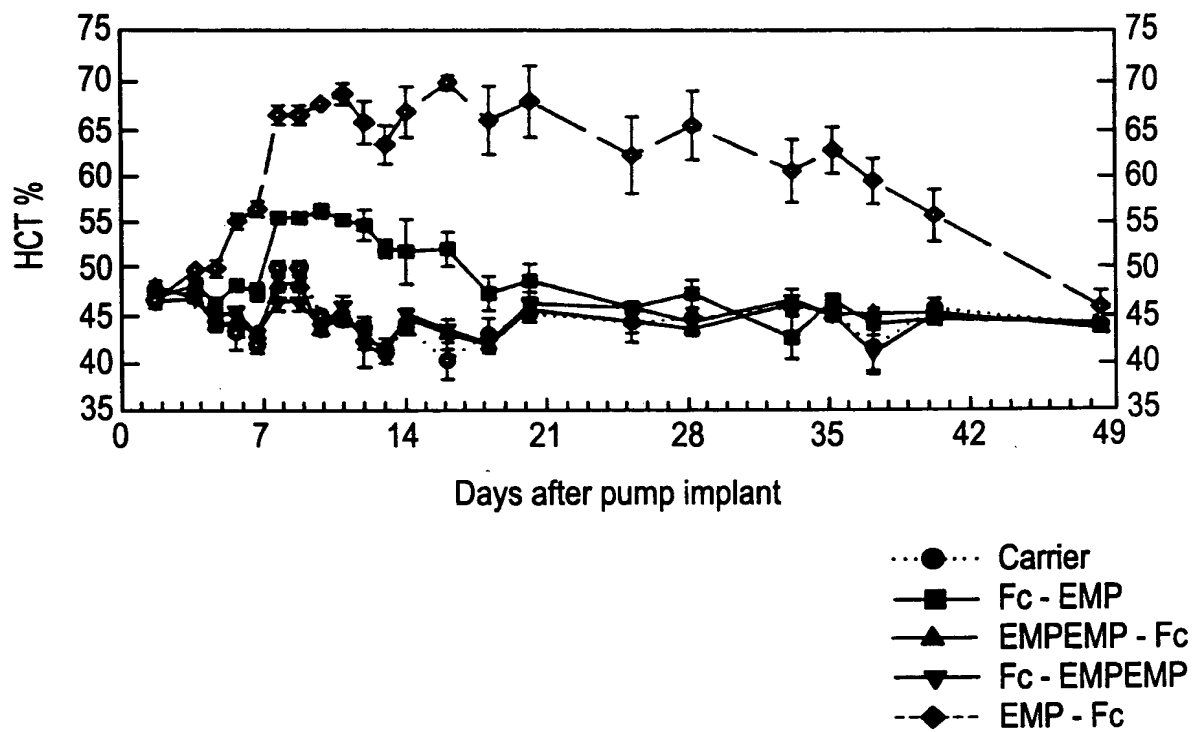


FIG. 19A

NdeI

```

1 CATATGGACAAAACTCACACATGTCCACCTTGTCCAGCTCCGGAACCTCTGGGGGGACCG
  -----+-----+-----+-----+-----+-----+-----+
a      M D K T H T C P P C P A P E L L G G P -
61 TCAGTCTTCCTCTTCCCCCAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAG
  -----+-----+-----+-----+-----+-----+-----+
a      S V F L F P P K P K D T L M I S R T P E -
121 GTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTAC
  -----+-----+-----+-----+-----+-----+-----+
a      V T C V V V D V S H E D P E V K F N W Y -
181 GTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGC
  -----+-----+-----+-----+-----+-----+-----+
a      V D G V E V H N A K T K P R E E Q Y N S -
241 ACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAG
  -----+-----+-----+-----+-----+-----+-----+
a      T Y R V V S V L T V L H Q D W L N G K E -
301 TACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAA
  -----+-----+-----+-----+-----+-----+-----+
a      Y K C K V S N K A L P A P I E K T I S K -
361 GCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTG
  -----+-----+-----+-----+-----+-----+-----+
a      A K G Q P R E P Q V Y T L P P S R D E L -
421 ACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCC
  -----+-----+-----+-----+-----+-----+-----+
a      T K N Q V S L T C L V K G F Y P S D I A -
481 GTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACACTACAAGACCACGCCTCCCGTGCTG
  -----+-----+-----+-----+-----+-----+-----+
a      V E W E S N G Q P E N N Y K T T P P V L -
541 GACTCCGACGGCTCCTTCTTCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAG
  -----+-----+-----+-----+-----+-----+-----+
a      D S D G S F F L Y S K L T V D K S R W Q -
600 CTGAGGCTGCCGAGGAAGAAGGAGATGTCGTTGAGTGGCACCTGTTCTCGTCCACCGTC
  -----+-----+-----+-----+-----+-----+-----+

```

FIG. 19B

```

601 CAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAG
-----+-----+-----+-----+-----+-----+ 660
GTCCCCTTGCAGAAGAGTACGAGGCACTACGTACTCCGAGACGTGTTGGTGATGTGCGTC

a      Q  G  N  V  F  S  C  S  V  M  H  E  A  L  H  N  H  Y  T  Q  -

661 AAGAGCCTCTCCCTGTCTCCGGGTAAAGGTGGAGGTGGTGGTGA CTTCCTGCCGCACTAC
-----+-----+-----+-----+-----+-----+ 720
TTCTCGGAGAGGGACAGAGGCCCATTTCCACCTCCACCACCACTGAAGGACGGCGTGATG

a      K  S  L  S  L  S  P  G  K  G  G  G  G  G  D  F  L  P  H  Y  -

                                     BamHI
                                     |
721 AAAAACACCTCTCTGGGTCACCGTCCGTAATGGATCC
-----+-----+-----+-----+-----+ 757
TTTTTGTGGAGAGACCCAGTGGCAGGCATTACCTAGG

a      K  N  T  S  L  G  H  R  P  *
```

FIG. 20A

NdeI
|

1 CATATGGACTTCCTGCCGCACTACAAAAACACCTCTCTGGGTACCGTCCGGGTGGAGGC 60
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
GTATACCTGAAGGACGGCGTGATGTTTTTGTGGAGAGACCCAGTGGCAGGCCCCACCTCCG

a M D F L P H Y K N T S L G H R P G G G -

61 GGTGGGGACAAACTCACACATGTCCACCTTGCCCAGCACCTGAACTCCTGGGGGGACCG 120
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CCACCCCTGTTTTGAGTGTGTACAGGTGGAACGGGTCGTGGACTTGAGGACCCCCCTGGC

a G G D K T H T C P P C P A P E L L G G P -

121 TCAGTTTTCCTCTTCCCCCAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAG 180
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
AGTCAAAAGGAGAAGGGGGGTTTTGGGTTCTGTGGGAGTACTAGAGGGCCTGGGGACTC

a S V F L F P P K P K D T L M I S R T P E -

181 GTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTAC 240
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CAGTGTAACGACCAACCTGCACTCGGTGCTTCTGGGACTCCAGTTCAAGTTGACCATG

a V T C V V V D V S H E D P E V K F N W Y -

241 GTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGC 300
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CACCTGCCGCACCTCCACGTATTACGGTTCTGTTTCGGGCGCCCTCCTCGTCATGTTGTGC

a V D G V E V H N A K T K P R E E Q Y N S -

301 ACGTACCGTGTGGTCAAGCTCCTCACCGTCTGCACCAGGACTGGCTGAATGGCAAGGAG 360
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TGCATGGCACACCAGTCGCAGGAGTGGCAGGACGTGGTCCTGACCGACTTACCGTTCTC

a T Y R V V S V L T V L H Q D W L N G K E -

361 TACAAGTGCAAGGTCTCCAACAAAGCCCTCCAGCCCCATCGAGAAAACCATCTCCAAA 420
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
ATGTTTACGTTCCAGAGGTTGTTTCGGGAGGGTGGGGGTAGCTCTTTTGGTAGAGGTTT

a Y K C K V S N K A L P A P I E K T I S K -

421 GCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTG 480
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CGGTTTCCCGTCGGGGCTCTTGGTGTCCACATGTGGGACGGGGGTAGGGCCCTACTCGAC

a A K G Q P R E P Q V Y T L P P S R D E L -

481 ACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCAGCGACATCGCC 540
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TGTTTCTTGGTCCAGTCGGACTGGACGGACAGTTTCCGAAGATAGGGTCGCTGTAGCGG

a T K N Q V S L T C L V K G F Y P S D I A -

541 GTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACACTACAAGACCACGCCTCCCGTGCTG 600
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CACCTCACCTCTCGTTACCCGTCGGCCTCTTGGTGTGTTCTGGTGGGAGGGGCACGAC

a V E W E S N G Q P E N N Y K T T P P V L -

FIG. 20B

```

601 GACTCCGACGGCTCCTTCTTCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAG
-----+-----+-----+-----+-----+-----+ 660
CTGAGGCTGCCGAGGAAGAAGGAGATGTCGTTGAGTGGCACCTGTTCTCGTCCACCGTC
a   D S D G S F F L Y S K L T V D K S R W Q .
661 CAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAG
-----+-----+-----+-----+-----+-----+ 720
GTCCCCCTTGAGAAGAGTACGAGGCACTACGTACTCCGAGACGTGTTGGTGATGTGCGTC
a   Q G N V F S C S V M H E A L H N H Y T Q .

                                     BamHI
                                     |
721 AAGAGCCTCTCCCTGTCTCCGGGTAAATAATGGATCCGCGG
-----+-----+-----+-----+-----+ 761
TTCTCGGAGAGGGACAGAGGCCCATTTATTACCTAGGCGCC
a   K S L S L S P G K *
```

FIG. 21A

NdeI
|
CATATGGACAAAACCTCACACATGTCCACCTTGTCCAGCTCCGGAACCTCCTGGGGGGACCG
1 -----+-----+-----+-----+-----+-----+-----+ 60
GTATACCTGTTTTGAGTGTGTACAGGTGGAACAGGTTCGAGGCCTTGAGGACCCCCCTGGC

a M D K T H T C P P C P A P E L L G G P -

TCAGTCTTCCTCTTCCCCCAAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAG
61 -----+-----+-----+-----+-----+-----+-----+ 120
AGTCAGAAGGAGAAGGGGGGTTTTGGGTTCTGTGGGAGTACTAGAGGGCCTGGGGACTC

a S V F L F P P K P K D T L M I S R T P E -

GTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTAC
121 -----+-----+-----+-----+-----+-----+-----+ 180
CAGTGTACGCACCACCACCTGCACTCGGTGCTTCTGGGACTCCAGTTCAAGTTGACCATG

a V T C V V V D V S H E D P E V K F N W Y -

GTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGC
181 -----+-----+-----+-----+-----+-----+-----+ 240
CACCTGCCGCACCTCCACGTATTACGGTTCTGTTTCGGCGCCCTCCTCGTCATGTTGTCG

a V D G V E V H N A K T K P R E E Q Y N S -

ACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAG
241 -----+-----+-----+-----+-----+-----+-----+ 300
TGCATGGCACACCAGTCGCAGGAGTGGCAGGACGTGGTCCTGACCGACTTACCGTTCCTC

a T Y R V V S V L T V L H Q D W L N G K E -

TACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAA
301 -----+-----+-----+-----+-----+-----+-----+ 360
ATGTTACAGTTCCAGAGGTTGTTTCGGGAGGGTCGGGGGTAGCTCTTTTGGTAGAGGTTT

a Y K C K V S N K A L P A P I E K T I S K -

GCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTG
361 -----+-----+-----+-----+-----+-----+-----+ 420
CGGTTTCCCGTCGGGGCTCTTGGTGTCCACATGTGGGACGGGGGTAGGGCCCTACTCGAC

a A K G Q P R E P Q V Y T L P P S R D E L -

ACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCC
421 -----+-----+-----+-----+-----+-----+-----+ 480
TGGTTCCTTGGTCCAGTCGGACTGGACGGACAGTTTCGAAGATAGGGTCGCTGTAGCGG

a T K N Q V S L T C L V K G F Y P S D I A -

GTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACATAAGACCACGCCTCCCGTGCTG
481 -----+-----+-----+-----+-----+-----+-----+ 540
CACCTCACCCCTCTCGTTACCCGTCGGCCTCTTGTGTATGTTCTGGTGGGAGGGCAGGAC

a V E W E S N G Q P E N N Y K T T P P V L -

GACTCCGACGGCTCCTTCTTCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAG
541 -----+-----+-----+-----+-----+-----+-----+ 600
CTGAGGCTGCCGAGGAAGAAGGAGATGTCGTTTCGAGTGGCACCTGTTCTCGTCCACCGTC

a D S D G S F F L Y S K L T V D K S R W Q -

FIG. 21B

```

601 CAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAG
-----+-----+-----+-----+-----+-----+ 660
GTCCCCTTGCAGAAGAGTACGAGGCACTACGTACTCCGAGACGTGTTGGTGATGTGCGTC

a   Q   G   N   V   F   S   C   S   V   M   H   E   A   L   H   N   H   Y   T   Q   -

661 AAGAGCCTCTCCCTGTCTCCGGGTAAAGGTGGAGGTGGTGGTTTCGAATGGACCCCGGGT
-----+-----+-----+-----+-----+-----+ 720
TTCTCGGAGAGGGACAGAGGCCCATTTCCACCTCCACCACCAAAGCTTACCTGGGGCCCA

a   K   S   L   S   L   S   P   G   K   G   G   G   G   G   F   E   W   T   P   G   -

                                     BamHI
                                     |
721 TACTGGCAGCCGTACGCTCTGCCGCTGTAATGGATCCCTCGAG
-----+-----+-----+-----+-----+ 763
ATGACCGTCGGCATGCGAGACGGCGACATTACCTAGGGAGCTC

a   Y   W   Q   P   Y   A   L   P   L   *

```

FIG. 22A

NdeI
|
1 CATATGTTTGAATGGACCCCGGGTTACTGGCAGCCGTACGCTCTGCCGCTGGGTGGAGGC 60
-----+-----+-----+-----+-----+-----+-----+-----+
GTATAACAAGCTTACCTGGGGCCCAATGACCGTCGGCATGCGAGACGGCGACCCACCTCCG
a M F E W T P G Y W Q P Y A L P L G G G -
61 GGTGGGGACAAAACCTCACACATGTCCACCTTGCCAGCACCTGAACTCCTGGGGGGACCG 120
-----+-----+-----+-----+-----+-----+-----+-----+
CCACCCCTGTTTTGAGTGTGTACAGGTGGAACGGGTCGTGGACTTGAGGACCCCCCTGGC
a G G D K T H T C P P C P A P E L L G G P -
121 TCAGTTTTCTCTTCCCCCAAACCAAGGACACCCTCATGATCTCCCGGACCCCTGAG 180
-----+-----+-----+-----+-----+-----+-----+-----+
AGTCAAAAGGAGAAGGGGGGTTTTGGGTTCCTGTGGGAGTACTAGAGGGCCTGGGGACTC
a S V F L F P P K P K D T L M I S R T P E -
181 GTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTAC 240
-----+-----+-----+-----+-----+-----+-----+-----+
CAGTGACGCACCACCACCTGCACTCGGTGCTTCTGGGACTCCAGTTCAAGTTGACCATG
a V T C V V V D V S H E D P E V K F N W Y -
241 GTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGC 300
-----+-----+-----+-----+-----+-----+-----+-----+
CACCTGCCGCACCTCCACGTATTACGGTTCTGTTTCGGCGCCCTCCTCGTCATGTTGTGC
a V D G V E V H N A K T K P R E E Q Y N S -
301 ACGTACCGTGTGGTCAGCGTCTCACCCTGTCACCAGGACTGGCTGAATGGCAAGGAG 360
-----+-----+-----+-----+-----+-----+-----+-----+
TGCATGGCACACCAGTCGCAGGAGTGGCAGGACGTGGTCCTGACCGACTTACCGTTCTCT
a T Y R V V S V L T V L H Q D W L N G K E -
361 TACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCATCGAGAAAACCATCTCCAAA 420
-----+-----+-----+-----+-----+-----+-----+-----+
ATGTTACGTTCCAGAGGTTGTTTCGGGAGGGTCGGGGGTAGCTCTTTTGGTAGAGGTTT
a Y K C K V S N K A L P A P I E K T I S K -
421 GCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTG 480
-----+-----+-----+-----+-----+-----+-----+-----+
CGGTTTCCCGTCGGGGCTCTTGGTGTCCACATGTGGGACGGGGGTAGGGCCCTACTCGAC
a A K G Q P R E P Q V Y T L P P S R D E L -
481 ACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCC 540
-----+-----+-----+-----+-----+-----+-----+-----+
TGGTTCTTGGTCCAGTCGGACTGGACGGACAGTTTCCGAAGATAGGGTCGCTGTAGCGG
a T K N Q V S L T C L V K G F Y P S D I A -
541 GTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACACTACAAGACCACGCCTCCCGTGCTG 600
-----+-----+-----+-----+-----+-----+-----+-----+
CACCTCACCTCTCGTTACCCGTCGGCCTCTTGTGTGATGTTCTGGTGGGAGGGGCACGAC
a V E W E S N G Q P E N N Y K T T P P V L -

FIG. 22B

```
601 GACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAG
-----+-----+-----+-----+-----+-----+ 660
CTGAGGCTGCCGAGGAAGAAGGAGATGTCGTTTCGAGTGGCACCTGTTCTCGTCCACCGTC
a   D S D G S F F L Y S K L T V D K S R W Q -

661 CAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAG
-----+-----+-----+-----+-----+ 720
GTCCCCTTGCAGAAGAGTACGAGGCACTACGTACTCCGAGACGTGTTGGTGATGTGCGTC
a   Q G N V F S C S V M H E A L H N H Y T Q -

                                     BamHI
                                     |
721 AAGAGCCTCTCCCTGTCTCCGGGTAAATAATGGATCC
-----+-----+-----+-----+ 757
TTCTCGGAGAGGGACAGAGGCCCATTTATTACCTAGG
a   K S L S L S P G K *
```

FIG. 23A

NdeI
|
CATATGGACAAAACCTCACACATGTCCACCGTGCCCAGCACCTGAACTCCTGGGGGGACCG
1 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 60
GTATACCTGTTTTGAGTGTGTACAGGTGGCACGGGTCTGTTGAGTGTGAGGACCCCCCTGGC

a M D K T H T C P P C P A P E L L G G P -

61 TCAGTTTTCTCTCTCCCCCAAACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 120
AGTCAAAGGAGAAGGGGGGTTTTGGGTTCCTGTGGGAGTACTAGAGGGCCTGGGGACTC

a S V F L F P P K P K D T L M I S R T P E -

121 GTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTAC
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 180
CAGTGTACGCACCACCACCTGCACTCGGTGCTTCTGGGACTCCAGTTCAAGTTGACCATG

a V T C V V V D V S H E D P E V K F N W Y -

181 GTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGC
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 240
CACCTGCCGCACCTCCACGTATTACGGTTCGTTCGCGGCCCTCCTCGTCATGTTGTCTG

a V D G V E V H N A K T K P R E E Q Y N S -

241 ACGTACCGTGTGGTCAGCGTCTCACCGTCTGCACCAGGACTGGCTGAATGGCAAGGAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 300
TGCATGGCACACCAGTCGCAGGAGTGGCAGGACGTGGTCCTGACCGACTTACCGTTCCTC

a T Y R V V S V L T V L H Q D W L N G K E -

301 TACAAGTGCAAGGTCTCCAACAAAGCCCTCCAGCCCCCATCGAGAAAACCATCTCCAAA
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 360
ATGTTACAGTTCCAGAGGTTGTTTCGGGAGGGTCTGGGGGTAGCTCTTTTGGTAGAGGTTT

a Y K C K V S N K A L P A P I E K T I S K -

361 GCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 420
CGGTTTCCCGTCTGGGGCTCTTGGTGTCCACATGTGGGACGGGGGTAGGGCCCTACTCGAC

a A K G Q P R E P Q V Y T L P P S R D E L -

421 ACCAAGAACCAGGTGACCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCC
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 480
TGGTTCTTGGTCCAGTCCGACTGGACGGACAGTTTCCGAAGATAGGGTCGCTGTAGCGG

a T K N Q V S L T C L V K G F Y P S D I A -

481 GTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACACTACAAGACCACGCTCCCGTGTCTG
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 540
CACCTCACCTCTCGTTACCGTCTGGCCTCTTGTGATGTTCTGGTGGGAGGGCAGCAG

a V E W E S N G Q P E N N Y K T T P P V L -

541 GACTCCGACGGCTCCTTCTTCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAG
-----+-----+-----+-----+-----+-----+-----+-----+-----+ 600
CTGAGGCTGCCGAGGAAGAAGGAGATGTCGTTTCGAGTGGCACCTGTTCTCGTCCACCGTC

a D S D G S F F L Y S K L T V D K S R W Q -

a I H V M W E W E C F E R L *

FIG. 24A

NdeI
|
1 CATATGGTTGAACCGAACTGTGACATCCATGTTATGTGGGAATGGGAATGTTTTGAACGT 60
-----+-----+-----+-----+-----+-----+-----+
GTATACCAACTTGGCTTGACACTGTAGGTACAATACACCCTTACCCTTACAAAACCTTGCA
a M V E P N C D I H V M W E W E C F E R -
61 CTGGGTGGTGGTGGTGGTGACAAAACACACATGTCCACCGTGCCCAGCACCTGAACTC 120
-----+-----+-----+-----+-----+-----+-----+
GACCCACCACCACCACCACTGTTTTGAGTGTGTACAGGTGGCAGGGTCGTGGACTTGAG
a L G G G G G D K T H T C P P C P A P E L -
121 CTGGGGGGACCGTCAGTTTTCTCTTCCCCCAAAACCAAGGACACCCTCATGATCTCC 180
-----+-----+-----+-----+-----+-----+-----+
GACCCCCCTGGCAGTCAAAGGAGAAGGGGGTTTTGGGTTCTGTGGGAGTACTAGAGG
a L G G P S V F L F P P K P K D T L M I S -
181 CGGACCCCTGAGGTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAG 240
-----+-----+-----+-----+-----+-----+-----+
GCCTGGGGACTCCAGTGTACGCACCACCACCTGCACTCGGTGCTTCTGGGACTCCAGTTC
a R T P E V T C V V V D V S H E D P E V K -
241 TTCAACTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAG 300
-----+-----+-----+-----+-----+-----+-----+
AAGTTGACCATGCACCTGCCGCACCTCCACGTATTACGGTTCTGTTTCGGCGCCCTCCTC
a F N W Y V D G V E V H N A K T K P R E E -
301 CAGTACAACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTG 360
-----+-----+-----+-----+-----+-----+-----+
GTCATGTTGTCGTGCATGGCACACCAGTCGCAGGAGTGGCAGGACGTGGTCTCTGACCGAC
a Q Y N S T Y R V V S V L T V L H Q D W L -
361 AATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCCATCGAGAAA 420
-----+-----+-----+-----+-----+-----+-----+
TTACCGTTCCTCATGTTACGTTCCAGAGGTTGTTTCGGGAGGGTCGGGGGTAGCTCTTT
a N G K E Y K C K V S N K A L P A P I E K -
421 ACCATCTCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCC 480
-----+-----+-----+-----+-----+-----+-----+
TGGTAGAGGTTTCGGTTTCCCGTCGGGGCTCTTGGTGTCCACATGTGGGACGGGGGTAGG
a T I S K A K G Q P R E P Q V Y T L P P S -
481 CGGGATGAGCTGACCAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCC 540
-----+-----+-----+-----+-----+-----+-----+
GCCCTACTCGACTGGTTCTTGGTCCAGTCGGACTGGACGGACCAGTTTCCGAAGATAGGG
a R D E L T K N Q V S L T C L V K G F Y P -
541 AGCGACATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACACTACAAGACCACG 600
-----+-----+-----+-----+-----+-----+-----+
TCGCTGTAGCGGCACCTCACCTCTCGTTACCCGTCGGCCTCTTGTGATGTTCTGGTGC
a S D I A V E W E S N G Q P E N N Y K T T -

FIG. 24B

```

601 CCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGGACAAG 660
    -----+-----+-----+-----+-----+-----+
    GGAGGGCACGACCTGAGGCTGCCGAGGAAGAAGGAGATGTCGTTTCGAGTGGCACCTGTTC
a      P P V L D S D G S F F L Y S K L T V D K -
    AGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAAC
661 -----+-----+-----+-----+-----+-----+ 720
    TCGTCCACCGTCGTCCCCTTGCAGAAGAGTACGAGGCACTACGTACTCCGAGACGTGTTG
a      S R W Q Q G N V F S C S V M H E A L H N -
                                     BamHI
                                     |
721 CACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATAACTCGAGGATCC 773
    -----+-----+-----+-----+-----+-----+
    GTGATGTGCGTCTTCTCGGAGAGGGACAGAGGCCCATTTATTGAGCTCCTAGG
a      H Y T Q K S L S L S P G K *
```

FIG. 25A

NdeI
|
1 CATATGGACAAAACCTCACACATGTCCACCTTGTCAGCTCCGGAACCTCCTGGGGGGACCG 60
-----+-----+-----+-----+-----+-----+-----+
GTATACCTGTTTTGAGTGTGTACAGGTGGAACAGGTCGAGGCCTTGAGGACCCCCCTGGC

a M D K T H T C P P C P A P E L L G G P -

61 TCAGTCTTCCTCTTCCCCCAAACCAAGGACACCCTCATGATCTCCCGGACCCCTGAG 120
-----+-----+-----+-----+-----+-----+-----+
AGTCAGAAGGAGAAGGGGGGTTTTGGGTTCTGTGGGAGTACTAGAGGGCCTGGGGACTC

a S V F L F P P K P K D T L M I S R T P E -

121 GTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTAC 180
-----+-----+-----+-----+-----+-----+-----+
CAGTGACGCACCACCACCTGCACTCGGTGCTTCTGGGACTCCAGTTCAAGTTGACCATG

a V T C V V V D V S H E D P E V K F N W Y -

181 GTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGC 240
-----+-----+-----+-----+-----+-----+-----+
CACCTGCCGCACCTCCACGTATTACGGTTCTGTTTCGGCGCCCTCCTCGTCATGTTGTGC

a V D G V E V H N A K T K P R E E Q Y N S -

241 ACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAG 300
-----+-----+-----+-----+-----+-----+-----+
TGCATGGCACACCAGTCGCAGGAGTGGCAGGACGTGGTCTGACCGACTTACCGTTCTCTC

a T Y R V V S V L T V L H Q D W L N G K E -

301 TACAAGTGCAAGGTCTCCAACAAAGCCCTCCAGCCCCCATCGAGAAAACCATCTCCAAA 360
-----+-----+-----+-----+-----+-----+-----+
ATGTTACAGTTCCAGAGGTTGTTTCGGGAGGGTCGGGGGTAGCTCTTTTGGTAGAGGTTT

a Y K C K V S N K A L P A P I E K T I S K -

361 GCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTG 420
-----+-----+-----+-----+-----+-----+-----+
CGGTTTCCCCTCGGGGCTCTTGGTGTCCACATGTGGGACGGGGGTAGGGCCCTACTCGAC

a A K G Q P R E P Q V Y T L P P S R D E L -

421 ACCAAGAACCAGGTGAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCAGCGACATCGCC 480
-----+-----+-----+-----+-----+-----+-----+
TGGTTCTTGGTCCAGTCGGACTGGACGGACCAGTTTCCGAAGATAGGGTCGCTGTAGCGG

a T K N Q V S L T C L V K G F Y P S D I A -

481 GTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACCTACAAGACCACGCCTCCCGTGCTG 540
-----+-----+-----+-----+-----+-----+-----+
CACCTCACCCCTCTCGTTACCCGTCGGCCTCTTGTGATGTTCTGGTGGGAGGGCAGGAC

a V E W E S N G Q P E N N Y K T T P P V L -

541 GACTCCGACGGCTCCTTCTCTCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAG 600
-----+-----+-----+-----+-----+-----+-----+
CTGAGGCTGCCGAGGAAGAAGGAGATGTCGTTTCGAGTGGCACCTGTTCTCGTCCACCGTC

a D S D G S F F L Y S K L T V D K S R W Q -

FIG. 25B

```

601 CAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTACACGCAG 660
    .....+.....+.....+.....+.....+.....+.....+.....+.....+
    GTCCCCTTGCAGAAGAGTACGAGGCACTACGTACTCCGAGACGTGTTGGTGATGTGCGTC

a      Q  G  N  V  F  S  C  S  V  M  H  E  A  L  H  N  H  Y  T  Q  -

    AAGAGCCTCTCCCTGTCTCCGGGTAAAGGTGGAGGTGGTGGTTGCACCACCACTGGGGT
661 .....+.....+.....+.....+.....+.....+.....+.....+.....+ 720
    TTCTCGGAGAGGGACAGAGGCCCATTTCCACCTCCACCACCAACGTGGTGGGTGACCCCA

A      K  S  L  S  L  S  P  G  K  G  G  G  G  G  C  T  T  H  W  G  -

          BamHI
          |
    TTCACCCTGTGCTAATGGATCCCTCGAG
721 .....+.....+.....+.....+.....+.....+.....+.....+.....+ 748
    AAGTGGGACACGATTACCTAGGGAGCTC

a      F  T  L  C  *
```

FIG. 26A

NdeI
 |
 1 CATATGTGCACCACCCACTGGGGTTTCACCCTGTGCGGTGGAGGCGGTGGGGACAAAGGT 60
 -----+-----+-----+-----+-----+-----+
 GTATACACGTGGTGGGTGACCCCAAAGTGGGACACGCCACCTCCGCCACCCCTGTTTCCA
 a M C T T H W G F T L C G G G G G D K G -
 61 GGAGGCGGTGGGGACAAACTCACACATGTCCACCTTGCCCAGCACCTGAACTCCTGGGG 120
 -----+-----+-----+-----+-----+-----+
 CCTCCGCCACCCCTGTTTGTAGTGTGTACAGGTGGAACGGGTCGTGGACTTGAGGACCCC
 a G G G G D K T H T C P P C P A P E L L G -
 121 GGACCGTCAGTTTTCTCTTCCCCCAAACCCAAAGGACACCCTCATGATCTCCCGGACC 180
 -----+-----+-----+-----+-----+-----+
 CCTGGCAGTCAAAGGAGAAAGGGGGGTTTTGGGTTCCTGTGGGAGTACTAGAGGGCCTGG
 a G P S V F L F P P K P K D T L M I S R T -
 181 CCTGAGGTCACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCTGAGGTCAAGTTCAAC 240
 -----+-----+-----+-----+-----+-----+
 GGACTCCAGTGTACGCACCACCACCTGCACTCGGTGCTTCTGGGACTCCAGTTCAAGTTG
 a P E V T C V V V D V S H E D P E V K F N -
 241 TGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGAGGAGCAGTAC 300
 -----+-----+-----+-----+-----+-----+
 ACCATGCACCTGCCGCACCTCCACGTATTACGGTTCTGTTTCGGCGCCCTCCTCGTCATG
 a W Y V D G V E V H N A K T K P R E E Q Y -
 301 AACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACTGGCTGAATGGC 360
 -----+-----+-----+-----+-----+-----+
 TTGTCGTGCATGGCACACCAGTCGCAGGAGTGGCAGGACGTGGTCTGACCGACTTACCG
 a N S T Y R V V S V L T V L H Q D W L N G -
 361 AAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCAGCCCCATCGAGAAAACCATC 420
 -----+-----+-----+-----+-----+-----+
 TTCCTCATGTTACGTTCCAGAGGTTGTTTCGGGAGGGTCGGGGGTAGCTCTTTTGGTAG
 a K E Y K C K V S N K A L P A P I E K T I -
 421 TCCAAAGCCAAAGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCCCCGGGAT 480
 -----+-----+-----+-----+-----+-----+
 AGGTTTTCGGTTTCCCGTCGGGGCTCTTGGTGTCCACATGTGGGACGGGGGTAGGGCCCTA
 a S K A K G Q P R E P Q V Y T L P P S R D -
 481 GAGCTGACCAAGAACCAGGTGACCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGAC 540
 -----+-----+-----+-----+-----+-----+
 CTCGACTGGTTCTTGGTCCAGTCGGACTGGACGGACCAGTTTCCGAAGATAGGGTCGCTG
 a E L T K N Q V S L T C L V K G F Y P S D -
 541 ATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGAGAACTACAAGACCACGCCTCCC 600
 -----+-----+-----+-----+-----+-----+
 TAGCGGCACCTCACCTCTCGTTACCCGTCGGCCTCTTGTGATGTTCTGGTGCAGGAGG
 a I A V E W E S N G Q P E N N Y K T T P P -

FIG. 26B

```

601 GTGCTGGACTCCGACGGCTCCTTCTTCTCTACAGCAAGCTCACCGTGGACAAGAGCAGG 660
-----+-----+-----+-----+-----+
CACGACCTGAGGCTGCCGAGGAAGAAGGAGATGTCGTTTCGAGTGGCACCTGTTCTCGTCC
a   V L D S D G S F F L Y S K L T V D K S R -

661 TGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCATGAGGCTCTGCACAACCACTAC 720
-----+-----+-----+-----+-----+
ACCGTCGTCCCCCTTGCAGAAGAGTACGAGGCACTACGTACTCCGAGACGTGTTGGTGATG
a   W Q Q G N V F S C S V M H E A L H N H Y -

                                     BamHI
                                     |
721 ACGCAGAAGAGCCTCTCCCTGTCTCCGGGTAAATAATGGATCC 763
-----+-----+-----+-----+
TGCGTCTTCTCGGAGAGGGACAGAGGCCCATTTATTACCTAGG
a   T Q K S L S L S P G K *
```